RSCP 21(2) 2016

Calcificación coronaria y depresión: un modelo de regresión logística.

Coronary calcification and depression: a logistic regression model.

Mariana Suárez-Bagnasco

Roberto Suárez-Antola

Resumen:
Este trabajo se divide en dos partes. En la primera parte estudiamos si los sujetos que presentan simultáneamente depresión estado y depresión rasgo tienen puntajes de calcio coronario más elevados que los sujetos con depresión estado solamente. En la segunda parte construimos un modelo de regresión logística multivariable y proponemos un predictor para la presencia de calcificación severa, basado en ese modelo.
En la primera parte, participaron sujetos entre 40 y 60 años, sin antecedentes coronarios, a los que se les hizo una angiotomografía cardíaca indicada por sus médicos tratantes. La depresión fue evaluada utilizando el Inventario de Depresión Estado/Rasgo de Spielberg. Hallamos asociaciones positivas entre cada una de las puntuaciones de las escalas y sub-escalas del test y los puntajes de calcio: eutimia estado 0,345, distimia estado 0,322, depresión estado 0,401, eutimia rasgo 0,376, distimia rasgo 0,325, depresión rasgo 0,407.
En la segunda parte, mediante un modelo de regresión logística multivariable calculamos la probabilidad de presencia de un puntaje de calcio mayor a 400 UH en función de un conjunto de variables predictoras (depresión estado, depresión rasgo, hipertensión arterial, dislipemia y edad). El modelo original fue simplificado, resultando un modelo final compuesto por un factor de riesgo cardiovascular tradicional (dislipemia), un factor de riesgo psicológico (depresión rasgo) y la edad como factor epidemiológico de riesgo. Una vez obtenidos los parámetros del modelo exploramos la posibilidad de construir un predictor. El valor óptimo para la probabilidad de corte cuando los pesos asignados a la sensibilidad y a la especificidad se tomaron iguales, se encontró comprendido entre 0,3 y 0,4. El área bajo la curva ROC resultó significativa a nivel $p<0,05$ con valor 0,74. La construcción de modelos como el que motiva la presente investigación, con un componente psicológico que no había sido incorporado en modelos previos, puede contribuir a la tarea de identificar nuevas interacciones relevantes desde una perspectiva epidemiológica.
Palabras clave: calcificación coronaria, modelo de regresión logística multivariable, depresión, dislipemia, edad.

Abstract
This research has two parts. In the first one we study if subjects that present simultaneously trait depression and state depression have a higher coronary calcium score than the subjects with state depression only. In the second part we construct a multivariable logistic regression model and propose a predictor of the presence of severe coronary calcification, based in the model.
In the first study participated subjects between 40 and 60 years old, without history of coronary disease, but that underwent a cardiac angiotomography by indication of their physicians. Depression was assessed by the Spielberg´s Inventary of Trait/ State Depression. We found positive associations between each one of the scores in the scales and sub-scales of the test and the calcium scores: state eutimia 0,345, state distimia 0,322, state depression 0,401, trait eutimia 0,376, trait distimia 0,325, trait depression 0,407.
In the second part, using a multivariable logistic regression model we estimated the probability of a coronary calcification with score greater than 400 UH, as a function of a set of prediction variables (state depression, trait depression, arterial hypertension, dislipemia and age). The original model was progressively simplified, until a final model was obtained with a traditional cardiovascular risk factor (dislipemia), a psychological risk factor (trait depression) and an epidemiological risk factor (age). Once the parameters of the model were obtained, we explored the possibility of constructing a predictor. When the weights assigned to the sensibility and to the specificity, were taken equal, the optimal value of the outage probability was comprised between 0.3 and 0.4. The area under the ROC curve was significant to level $p<0.05$ with value 0.74. The construction of models like the one that motivates the present research, with a psychological component not considered in previous models, could contribute to the task of identifying new relevant interactions from an epidemiological point of view.
Key words: coronary calcification, multivariable logistic regression model, depression, dislipemia, age.

Introducción

La enfermedad ateroesclerótica de las arterias coronarias es una de las principales causas de mortalidad en países occidentales. Es una enfermedad multifactorial que involucra entre sus principales mecanismos la disfunción endote-

lial vascular y la inflamación vascular. Suele desarrollarse y evolucionar de manera asintomática pudiéndose manifestar clínicamente a través del síndrome coronario agudo, infarto agudo de miocardio, angina inestable o muerte súbita [1-5].

Sin embargo, la enfermedad coronaria puede detectarse antes de que se manifieste clínicamente cuantificando el depósito de calcio de las arterias coronarias ("score cálcico") dado que el puntaje de calcio es considerado un buen predictor de eventos coronarios agudos. Por ejemplo, si comparamos sujetos con puntaje de calcio 0 y sujetos con puntaje de calcio entre 1 y 100 Unidades Hounsfield (UH) o sujetos con más de 300 UH, los segundos presentan entre tres y cuatro veces más riesgo de eventos coronarios agudos que los primeros y los terceros presentan entre siete y 10 veces más riesgo que los primeros.

La cuantificación del calcio de las arterias coronarias puede realizarse utilizando un tomógrafo computado multicorte, no requiere la inyección de contraste iodado y utiliza dosis efectivas relativamente bajas de radiación gamma (0.5 a 1.5 mSv) en comparación con otros exámenes mediante tomografía computada.
Generalmente, se considera que los puntajes de calcio entre 1 y 100 UH indican la presencia de calcificación de grado leve, los puntajes entre 101 y 400 UH indican la presencia de calcificación de grado moderado y los puntajes de calcio a partir de 401 UH son un indicador de presencia de calcificación severa [6-9].
Algunos estudios han explorado la posible relación entre factores de riesgo psicosociales como la depresión, la ansiedad y el estrés, y el puntaje de calcio coronario. Si bien los trabajos que han estudiado la asociación entre el puntaje de calcio y la depresión evaluada en una sola instancia aportan resultados contradictorios, los dos trabajos que estudiaron la relación entre la presencia de episodios repetidos de depresión y el puntaje de calcio aportan evidencias acerca de una asociación entre dichas variables. Un estudio evaluó la depresión en tres instancias a lo largo de diez años y otro estudio evalúo la depresión en cinco instancias a lo largo de cinco años. Los autores que evaluaron la depresión anualmente a lo largo de cinco años concluyeron que los participantes con al menos tres episodios depresivos a lo largo de cinco años tenían dos veces más probabilidad de presentar un puntaje de calcio coronario mayor a 10 UH que las personas que no presentaron episodios depresivos [10-14].

Desde un punto de vista clínico, la presencia de frecuentes episodios de depresión podría sugerir la existencia de una predisposición relativamente estable, una disposición general de afectividad que hace más probable presentar no solamente depresión estado sino también varios episodios de depresión estado en el trascurso del tiempo, lo que se denomina depresión rasgo.

En la primera parte de este trabajo estudiamos si los sujetos que presentan simultáneamente depresión estado y depresión rasgo presentan puntajes de calcio más elevados que los sujetos con depresión estado solamente. En la segunda parte proponemos un predictor para la presencia de calcificación severa, basado en un modelo de regresión logística multivariable.

Primera parte

Materiales y métodos

Para este trabajo se utilizó una base de datos que incluía 200 sujetos de una población urbana, entre 40 y 60 años, sin antecedentes coronarios, que se realizaron un angiotomografía cardíaca derivados por sus médicos tratantes, en un mismo servicio hospitalario.
A todos los sujetos les realizaron una angiotomografía cardíaca utilizando un Tomógrafo Axial Computado Multicorte 64 hileras de detectores (Aquilion 64, Toshiba, Japón).
La presencia de factores de riesgo cardiovasculares tradicionales se recabaron a través de la anamnesis, considerándolos como variables dicotómicas.
La depresión fue evaluada utilizando el Inventario de Depresión Estado/Rasgo (IDER) diseñado por C. D. Spielberg, adaptado al castellano [15]. Está constituido por 20 ítems distribuidos en dos escalas: depresión estado y depresión rasgo. La depresión estado alude a una condición emocional o mental puntual, transitoria, que suele corresponderse con la presencia de un episodio depresivo mayor o un trastorno depresivo mayor. La depresión rasgo refiere a una predisposición relativamente estable a responder de una determinada manera a las situaciones particulares y puntua-

les, una disposición general de afectividad que aumenta la posibilidad de presentar depresión estado. Cada una de estas escalas, depresión estado y depresión rasgo, incluye a su vez dos sub-escalas: distimia estado, eutimia estado, distimia rasgo y eutimia rasgo, respectivamente. Todas las puntuaciones del IDER están diseñadas de modo que un mayor valor numérico corresponde a un indicio más fuerte de depresión. Así, pese al significado etimológico de la palabra eutimia, en el IDER una puntuación alta en esa sub-escala indica baja afectividad positiva, porque la eutimia refiere a la ausencia de afectos positivos. Por su parte la distimia se refiere a la presencia de afectos negativos. Adicionalmente, las puntuaciones obtenidas a partir de la administración de cada una de las escalas fueron consideradas como variables discretas tomando como punto de corte el percentil 75 [15].
Se excluyeron los sujetos que presentaban calcificaciones en la aorta y/o válvulas, y/o lesiones ateroscleróticas no cálcicas o mixtas [16], resultando una muestra de 90 sujetos, edad 48,18±9,22 años, 79 hombres y 18 mujeres[1].
Utilizando el paquete de software SPPS 12 se calcularon estadísticos descriptivos y se ensayaron correlaciones. Para ensayar correlaciones entre variables continuas se calcularon coeficientes r de Pearson y para ensayar correlaciones entre una variable continua y una variable dicotómica se calcularon coeficientes biseriales puntuales. Para estudiar si existían diferencias entre las medias de los puntajes de calcio se utilizó la prueba Kruskal-Wallis.

Resultados

De los 90 sujetos, 76 presentaban dislipemia, 72 hipertensión arterial, 40 obesidad, 40 antecedentes heredofamiliares, 15 tabaquismo y 10 diabetes.
Los puntajes de calcio coronario de los participantes oscilaron entre 10 y 1029 UH.

Los sujetos que presentaron simultáneamente depresión estado y depresión rasgo presentaron medias de puntajes de calcio mayores que los sujetos con depresión estado solamente o con depresión rasgo solamente, siendo 988,33 UH, 186,21 UH y 541,29 UH, respectivamente. Se hallaron diferencias estadísticamente significativas entre estas medias (0.335, p=0.0321).

Al ensayar correlaciones entre el puntaje de calcio coronario y los factores de riesgo cardiovasculares tradicionales por un lado y el puntaje de calcio coronario y datos sociodemográficos por otro lado, se hallaron correlaciones estadísticamente significativas entre el puntaje de calcio coronario y: a) dislipemia (0,338, p=0,024), b) hipertensión arterial (0,212, p=0,039), c) edad (0,356, p=0,028).

Al ensayar correlaciones estadísticas entre el puntaje de calcio coronario y los puntajes obtenidos a partir de la administración de la prueba psicológica, hallamos asociaciones positivas entre cada una de las puntuaciones de las escalas y subescalas de la prueba y los puntajes de calcio: eutimia estado 0,345, distimia estado 0,322, depresión estado 0,401, eutimia rasgo 0,376, distimia rasgo 0,325, depresión rasgo 0,407.

Conclusiones de la primera parte

A partir de los resultados obtenidos analizando la muestra, se puede concluir que la presencia simultánea de depresión estado y depresión rasgo se asoció a valores medios de puntajes de calcio más elevados que la presencia de depresión estado solo o depresión rasgo solamente.
Los resultados de esta investigación preliminar sugieren que cuando las personas presentan simultáneamente como estado y como rasgos afectos negativos (distimia) junto a la ausencia de afectos positivos (eutimia) como estado y como rasgo, es más probable que presenten mayores puntajes de calcio, y por ello mayor riesgo de presentar un evento coronario agudo.
Además, la media del puntaje de calcio de los sujetos con depresión rasgo fue mayor que la media de los sujetos con depresión estado solamente. Esto sugiere que sería necesario no solo identificar a los sujetos que pueden encontrarse deprimidos en el momento de la evaluación sino también sería necesario identificar aquellos que si bien en el momento de la evaluación pueden no presentan depresión estado, pueden presentar una predisposición que los hace vulnerables a presentar con frecuencia depresión estado.

[1] Dadas las características de la muestra, no se pudo tomar en cuenta el sexo.

Una posible explicación de las correlaciones positivas encontrados entre los puntajes de calcio y las puntuaciones en las escalas de depresión se podría basar en los ambientes fisiopatológicos comunes a la depresión y la aterosclerosis coronaria como son por ejemplo el predomino de la actividad simpática respecto a la actividad parasimpática, el incremento de los niveles de catecolaminas, y la liberación de citoquinas proinflamatorias [17-22]. La presencia de frecuentes episodios de depresión estado podría favorecer el incremento del grado de calcificación de las coronarias.
Si los resultados obtenidos en esta investigación piloto se confirmaran en trabajos posteriores, podría ser importante que los cardiólogos cuyos pacientes presenten calcificación en las coronarias evalúen la posibilidad de derivarlos al servicio de salud mental y los psiquiatras cuyos pacientes presenten depresión tengan en cuenta la posibilidad de derivar a estos pacientes a cardiología para evaluar el puntaje de calcio de sus coronarias.

Segunda parte

Dadas las correlaciones positivas halladas entre el puntaje de calcio y: a) hipertensión arterial, b) dislipemia, c) depresión (tanto estado como rasgo) y d) edad, en función de estos resultados, nos planteamos la siguiente pregunta ¿cómo se modifica el riesgo de presencia de calcificación severa (más de 400 UH) en un paciente en función de las variables depresión estado, depresión rasgo, hipertensión arterial, dislipemia y edad?
Para responder a esta pregunta ensayamos la posibilidad de representar la probabilidad de presencia de puntaje de calcio mayor a 400 UH mediante una regresión logística multivariable, de modo de describir la influencia de cada variable predictora (covariable del modelo) mediante un coeficiente de regresión correspondiente.
Los coeficientes de regresión se determinan ajustando el modelo matemático a los datos experimentales obtenidos en la población bajo estudio. Una vez hallados los coeficientes, el modelo suministra una fórmula para calcular la probabilidad de puntaje de calcio mayor a 400 UH en función de los valores que pueden tomar las covariables.

La elección del modelo se efectuó a partir de las siguientes consideraciones. Los modelos de regresión logística multivariable se han venido utilizando en epidemiología y clínica médica desde hace tiempo [21,23]. En cardiología, se aplican al menos desde comienzos de la década del sesenta del siglo pasado. Un ejemplo representativo es el trabajo de Truett, Cornfield y Kannel basado en Framingham [24]. Como consecuencia se cuenta con amplia experiencia sobre las fortalezas y las debilidades de estos modelos [25].
La regresión logística utilizada es del tipo denominado binomial. En base a la experiencia disponible en relación con la aplicación de modelos binomiales de regresión logística, cabe esperar que un modelo binomial permita obtener una estimación no sesgada de la probabilidad de presencia de puntaje de calcio mayor a 400 UH a partir de los valores de las covariables consideradas simultáneamente [26-28].
Por último, disponemos de un software adecuado y probado, las diferentes versiones del SPSS, para el desarrollo de este tipo de modelo estadístico no lineal.
Una vez obtenidos los parámetros del modelo exploramos la posibilidad de construir una herramienta para estimar la presencia de calcificación severa.

Métodos

De acuerdo con lo mencionado en la introducción y con lo descripto en la bibliografía [29-32] seleccionamos las variables hipertensión arterial, dislipemia, depresión estado, depresión rasgo y edad como covariables (variables predictoras), y puntaje de calcio mayor a 400 UH aterosclerótica como variable dependiente (variable predicha).
Las variables depresión estado, depresión rasgo, hipertensión arterial, dislipemia y puntaje de calcio mayor a 400 UH coronaria fueron consideradas variables cualitativas de nivel de medición nominal, ya que cada una implicaba dos opciones, 0: ausencia de la característica de estudio, 1: presencia de la característica de estudio. La edad se agrupó en dos categorías 40 a 50 (a la que se le asigna valor 0) y 51 en adelante (a la que se le asigna valor 1).

Para concretar la posibilidad de relacionar las variables estudiadas empleando un modelo de regresión logística binomial multivariable, de modo de poder utilizarlo para fines prácticos, se requiere:

1) Determinar las covariables que se considera deben aparecer en la fórmula que resume el modelo matemático de regresión logística.
2) Estimar los valores de los coeficientes de regresión que corresponden a cada una de las variables para la población considerada, a partir de los datos experimentales.

La primera tarea se llevó a cabo en tres pasos, siguiendo un método ya establecido [24, 26-28]. El primer paso consistió en la exploración de asociaciones bivariantes para identificar posibles asociaciones entre cada una de las covariables y la variable dependiente, con el fin de determinar las covariables que podrían aparecer en la fórmula. En particular se consideraron la posibilidad de hacer aparecer términos de interacción en el modelo (representados por productos de covariables) y la posibilidad de colinealidad entre alguna de las covariables. Fue realizado en la primera parte considerando como significativas las correlaciones a un nivel de significación usualmente utilizado en la literatura, $p<0,05$.
En un segundo paso se analizaron correlaciones y posibles efectos de confusión en las variables seleccionadas[2].
En un tercer paso, calculamos el número máximo de covariables que podíamos emplear en el modelo de regresión logística multivariable siguiendo el procedimiento establecido en [33].

La segunda tarea también se efectuó en etapas, siguiendo recomendaciones ya establecidas [24, 26-28].
Se comenzó por la construcción de una primera versión del modelo con las variables seleccionadas en función de los resultados obtenidos en la investigación previa:

$$\log_e\left(\frac{p}{1-p}\right) = B_0 + B_1 \times DE + B_2 \times DR + B_3 \times E + B_4 \times DLP + B_5 \times HTA$$

Esta fórmula relaciona la probabilidad de presencia de puntaje de calcio mayor a 400 UH p con las covariables depresión estado *DE*, depresión rasgo *DR*, edad *E*, dislipemia DLP e hipertensión arterial HTA.
Los coeficientes de regresión B_1, B_2, B_3, B_4, B_5 y el denominado coeficiente de intersección B_0 constituyen los parámetros de este modelo logístico.

A la izquierda de la igualdad aparece el logaritmo natural del cociente $O = \frac{p}{1-p}$ entre la probabilidad de que el puntaje de calcio mayor a 400 UH esté presente y la probabilidad de que no esté presente. Este cociente de probabilidades, generalizado para un evento dicotómico cualquiera, es lo que en inglés se denomina **odds,** en español se suele denominar oportunidad y aquí representamos mediante O. Entre la oportunidad (odds) y la probabilidad hay una relación monótona creciente: cuando aumenta la probabilidad aumenta la oportunidad y viceversa.

Si se conocen los parámetros del modelo, O al igual que p, quedan determinadas una vez que se fijan los valores de las covariables.
Cuando se fijan los valores de todas las covariables excepto una de ellas, digamos la dislipemia (DLP), y a $\log_e O(DLP=0)$ se le resta $\log_e O(DLP=1)$, se obtiene el logaritmo del cociente de oportunidades u oportunidad relativa (odds ratio) construido dividiendo la oportunidad cuando la dislipemia está presente entre la oportunidad cuando no lo está: $\log_e \frac{O(DLP=1)}{O(DLP=0)} = B_4$ De manera equivalente $\frac{O(DLP=1)}{O(DLP=0)} = e^{B_4}$.
Un razonamiento análogo se puede realizar para conectar los demás coeficientes de regresión B_1, B_2, B_3 y B_5 con las oportunidades relativas correspondientes al cambio de valor de 0 a 1 en cada una de las covariables *DE, DR, E* y HTA, mientras las demás permanecen fijas.

Cuando un coeficiente de regresión es positivo le corresponde una oportunidad relativa mayor que uno. En ese caso, cuando la covariable pasa de 0 a 1, aumenta la oportunidad de que el puntaje de calcio mayor a 400 UH esté presente. A su vez ese aumento será tanto mayor cuanto mayor sea el valor del coeficiente de regresión. Por ende será mayor el efecto de la presencia del factor estadístico de riesgo representado por esa covariable. Por el contrario, un coeficiente de regresión negativo corresponde a una oportunidad relativa menor que uno: cuanto mayor sea en valor absoluto ese coeficiente, tanto mayor será el efecto de disminución de la oportunidad de que el puntaje de calcio mayor a 400 UH esté presente cuando la variable pasa de 0 a 1. Por ende, será mayor el efecto de la presencia del

[2] Teniendo en cuenta que las correlaciones y posibles variables confundentes asociadas a las covariables del modelo han sido estudiadas en numerosos trabajos [16].

factor estadístico (de protección porque su presencia disminuye la probabilidad de presencia de puntaje de calcio mayor a 400 UH) representado por la covariable.

Introducimos la denominación de factor estadístico de riesgo/protección para distinguirlo de los factores de riesgo cardiovasculares tradicionales que son un caso particular de estos factores estadísticos[3].
Utilizamos el programa estadístico SPSS 12 para todos los procesos de cálculo que se describen a continuación.
Se realizó una primera determinación de los coeficientes de regresión e intersección. El software utilizado estima esos coeficientes aplicando un método de máxima verosimilitud (busca combinaciones de parámetros de regresión que maximicen la probabilidad de obtener la muestra observada). Se analizaron los coeficientes de regresión de este primer modelo logístico para cada una de las variables, teniendo en cuenta las incertidumbres en su estimación, y se redujo el número de variables aplicando el criterio de parsimonia (navaja de Occam).
Para analizar la posibilidad de reducir el número de variables se ensayaron distintos modelos de regresión, relacionando el puntaje de calcio mayor a 400 UH con todas las combinaciones posibles de las covariables (factores de riesgo cardiovascular tradicionales, factores psicológicos, factores epidemiológicos) tomadas aisladamente y en conjunto.
De esta forma se eliminaron variables hasta obtener un modelo final que involucra un factor de riesgo cardiovascular tradicional, un factor psicológico, un factor epidemiológico. El umbral de 400 UH se adoptó siguiendo un criterio ampliamente utilizado en la práctica [6-9].

A partir de la fórmula que relaciona la probabilidad de presencia de puntaje de calcio mayor a 400 UH con las tres covariables seleccionadas se construyó un predictor.

Para construir el predictor se determinó un valor de corte de la probabilidad de presencia de calcificación severa tal que cuando esa probabilidad se encuentra por encima del valor de corte, se asume la presencia de un puntaje de calcio mayor a 400 UH.

Para elegir un valor de corte se determinaron la sensibilidad y la especificidad del predictor para diez valores de la probabilidad de corte elegidos a intervalos iguales. Se determinó un punto de corte mediante un criterio de optimización basado en un promedio ponderado de la sensibilidad y la especificidad.
Luego, construimos la curva ROC (Receiver Operating Characteristic) graficando la sensibilidad contra 1-especificidad para cada uno de los diferentes valores de la probabilidad de corte y analizamos el área bajo la curva considerada como una medida global de la validez del predictor.

Por último, evaluamos la bondad del ajuste del modelo de regresión logística (con tres factores) mediante el método que emplea la desviación [27] como medida de la falta de ajuste del modelo a los datos experimentales.
La desviación se define a partir de la verosimilitud L como:

$$D(\{B_k\}) = -2 \times \log_e [L\{B_k\}]$$

Es función de los valores que se asignen a cada uno de los parámetros del modelo, representados en conjunto mediante $\{B_k\}$.
Puesto que la verosimilitud $L\{B_k\}$ es la probabilidad de obtener la muestra observada dado un conjunto $\{B_k\}$ de valores asumidos para los parámetros del modelo, en principio la desviación, como función de los valores $\{B_k\}$ puede tomar cualquier valor no negativo, a partir de 0 (cuando la verosimilitud vale 1 y el ajuste del modelo a los datos es perfecto) creciendo sin límite a medida que la verosimilitud se aproxima a cero (y el ajuste del modelo a los datos es cada vez peor).
Calculamos las desviaciones para el modelo ajustado (con los valores hallados para el coeficiente de intersección B_0 y los coeficientes de regresión B_k correspondientes a cada covariable por separado y en conjunto) y para el modelo nulo (con el coeficiente de intersección hallado y todos los coeficientes de regresión nulos). Luego comparamos las desviaciones.

[3] Introducimos la denominación de factor estadístico de riesgo/protección para distinguirlo de los factores de riesgo cardiovasculares tradicionales que son un caso particular de estos factores estadísticos.

Resultados

La estimación de los parámetros correspondientes al modelo preliminar arrojó los siguientes resultados en la tabla 1.

						Intervalo de confianza al 95% para Exp(B)	
	B	Error típico	Wald	Sig.	Exp(B)	Límite inferior	Límite superior
Intersección	-3,099	0,712	18,962	0,000			
DE	0,368	0,491	0,56	0,454	1,444	0,551	3,783
DR	0,482	0,508	0,899	0,343	1,619	0,598	4,381
E	1,459	0,447	10,648	0,001	4,302	1,791	10,334
DLP	1,555	0,413	14,17	0,000	4,733	2,107	10,633
HTA	0,05	0,405	0,015	0,903	1,051	0,475	2,324

Tabla 1. Estimación de los parámetros correspondientes al modelo preliminar. DE: depresión estado, DR: depresión rasgo, E: edad, DLP: dislipemia, HTA: hipertensión arterial.

Obsérvese que la variable hipertensión arterial presenta un Exp(B) de 1,051 con un intervalo de confianza de la estimación comprendido entre 0,475 y 2,324. El valor estimado para Exp(B) y su intervalo de confianza, así como una significación cercana a uno y un estadístico de Wald muy pequeño, sugieren que esta variable de riesgo cardiovascular tradicional puede ser excluida en el marco de este modelo.

Para el caso de las variables psicológicas, sus Exp(B) estimados, 1,619 para la depresión rasgo y 1,444 para la depresión estado, implican una variación no despreciable en el odds para la presencia de puntaje de calcio mayor a 400 UH.
No obstante, cuando se examinan los correspondientes intervalos de confianza, se hallan límites inferiores 0,598 y 0.551. Esto constituye un problema en la medida en que implica que dentro del intervalo de confianza para Exp(B) aparecen valores superiores e inferiores a uno. Pero a diferencia de lo que se observó para la variable hipertensión, las distribuciones de las Exp(B) ya no quedan centradas en uno, sino que están centradas en valores mayores a uno. Por este motivo, en una primera instancia dejamos ambas variables psicológicas.

Repetimos el proceso, ahora con las covariables depresión estado, depresión rasgo, dislipemia y edad, y obtuvimos la siguiente estimación de los parámetros del modelo logístico, confirmando que la decisión de excluir la variable hipertensión arterial fue adecuada. Ver tabla 2

						Intervalo de confianza al 95% para Exp(B)	
	B	Error típico	Wald	Sig.	Exp(B)	Límite inferior	Límite superior
Intersección	-3,09	0,708	19,029	0,000			
DE	0,367	0,491	0,559	0,455	1,443	0,551	3,778
DR	0,489	0,504	0,943	0,332	1,631	0,607	4,381
E	1,46	0,447	10,648	0,001	4,306	1,791	10,349
DLP	1,575	0,378	17,403	0,000	4,832	2,305	10,128

Tabla 2. Estimación de los parámetros correspondientes al modelo con las variables: DE: depresión estado, DR: depresión rasgo, E: edad y DLP: dislipemia.

La fórmula para la probabilidad de presencia de puntaje de calcio mayor a 400 UH expresada mediante el logaritmo de la oportunidad de presencia de puntaje de calcio mayor a 400 UH queda entonces completamente especificada:

$$\log_e\left(\frac{p}{1-p}\right) = -3{,}090 + 0{,}367 \times DE + 0{,}489 \times DR + 1{,}460 \times E + 1{,}575 \times DLP$$

Despejando la probabilidad de presencia de puntaje de calcio mayor a 400 UH a partir de esta fórmula, se obtiene:

$$p = \frac{e^{-3{,}090+0{,}367\times DE+0{,}489\times DR+1{,}460\times E+1{,}575\times DLP}}{1+e^{-3{,}090+0{,}367\times DE+0{,}489\times DR+1{,}460\times E+1{,}575\times DLP}}$$

Los estadísticos para la depresión rasgo mejoraron desde el punto de vista de su aporte al modelo, mientras que no ocurre lo mismo para los estadísticos de la variable depresión estado.

Se observa que el problema planteado en relación con el intervalo de confianza para la Exp(B) de las variables psicológicas se mantiene. Nos ocupamos de esto eliminando una de las variables psicológicas. Eliminamos la variable depresión estado por el comportamiento de los estadísticos ya mencionados y por razones clínicas. La presencia de puntaje de calcio mayor a 400 UH implica un proceso evolutivo temporalmente extenso que cabe esperar asociado a un rasgo psicológico depresivo más que a un estado depresivo más delimitado en un marco temporal breve.

Repetimos el proceso, ahora con las covariables depresión rasgo, dislipemia y edad, y obtuvimos la siguiente estimación de los parámetros del modelo logístico, confirmando que la decisión de excluir la variable depresión estado fue adecuada. Ver tabla 3.

						Intervalo de confianza al 95% para Exp(B)	
	B	Error típico	Wald	Sig.	Exp(B)	Límite inferior	Límite superior
Intersección	-3,043	0,705	18,629	0,000			
DLP	1,585	0,377	17,65	0,000	4,878	2,329	10,218
E	1,455	0,446	10,628	0,001	4,286	1,787	10,283
DR	0,74	0,379	3,802	0,051	2,095	0,996	4,407

Tabla 3. Estimación de los parámetros correspondientes al modelo con las variables DR: depresión rasgo, E: edad y DLP: dislipemia.

Como se observa en la tabla, cuando la variable depresión rasgo se presenta como único factor psicológico, mejoran sus estadísticos.

Debemos destacar los siguientes cambios: Exp(B) pasa de 1,631 a 2,095, el estadístico de Wald pasa de 0,943 a 3,802, y el intervalo de confianza pasa de 0,607-4,481 a 0,996-4,407.
En particular hay un descenso en el límite superior y un ascenso en el límite inferior del intervalo. Este último es trascendente ya que es casi uno, con lo cual desaparece el problema de la dualidad del efecto sobre el odds de presencia de puntaje de calcio mayor a 400 UH que ya señalamos.

La fórmula para la probabilidad de presencia de puntaje de calcio mayor a 400 UH expresada mediante el logaritmo de la oportunidad queda entonces completamente especificada:

$$\log_e\left(\frac{p}{1-p}\right) = -3{,}043 + 0{,}740 \times DR + 1{,}455 \times E + 1{,}585 \times DLP$$

Despejando la probabilidad de presencia de puntaje de calcio mayor a 400 UH a partir de esta fórmula, se obtiene:

$$p = \frac{e^{-3,043+0,740\times DR+1,455\times E+1,585\times DLP}}{1+e^{-3,043+0,740\times DR+1,455\times E+1,585\times DLP}}$$

Dado que en todos los casos ensayados, tanto la edad como la dislipemia muestan un comportamiento dominante respecto a las otras variables, ensayamos eliminar la variable psciológica dejando solamente estas variables dominantes en el modelo de regresión.

Se obtuvieron los resultados que aparecen en la tabla 4.

						Intervalo de confianza al 95% para Exp(B)	
	B	Error típico	Wald	Sig.	Exp(B)	Límite inferior	Límite superior
Intersección	-2,567	0,641	16,049	0,000			
E	1,448	0,442	10,712	0,001	4,255	1,788	10,129
DLP	1,57	0,372	17,838	0,000	4,808	2,32	9,964

Tabla 4. Estimación de los parámetros correspondientes al modelo con las variables E: edad y DLP: dislipemia.

Como se observa en la tabla cuando se elimina del modelo la variable depresión rasgo hay un cambio considerable en el parámetro de intersección: B pasa de -3,043 en el modelo anterior con la variable psicológica a -2,567 en el modelo sin la variable psicológica, y el estadístico de Wald pasa de 18,629 a 16,049.

Cuando se estima la probabilidad de presencia de puntaje de calcio mayor a 400 UH cuando las covariables toman el valor 0 (ausencia) se obtiene para el modelo con y sin variable psicológica, respectivamente:

$$p(DLP=0,E=0,DR=0)=\frac{e^{-3,043}}{1+e^{-3,043}}=0,045 \qquad p(DLP=0,E=0)=\frac{e^{-2,567}}{1+e^{-2,567}}=0,071$$

Comparando ambas estimaciones de la probabilidad de presencia de puntaje de calcio mayor a 400 UH se observa que cuando la variable psicológica no está presente esa probabilidad es 1,6 veces mayor que el valor que toma en presencia de la variable psicológica en el marco del modelo.

Cuando todos los coeficientes de regresión son positivos el efecto de cambiar el valor de una o más covariables de 0 a 1, conduce a un aumento de la probabilidad de presencia de puntaje de calcio mayor a 400 UH.
En este caso, como regla práctica, el parámetro de intersección (B_0) debe ser, en valor absoluto, mayor a 3, para que la probabilidad de presencia de puntaje de calcio mayor a 400 UH en ausencia de los factores que predisponen a ella sea menor a 0,05.
Utilizamos este regla porque un valor menor o igual a 0,05 es uno de los más aceptados en estadística.
No utilizamos el 0,01 debido a que no sería una exigencia acorde con las incertidumbres del modelo incluyendo la posible existencia de otras variables predictoras que no se contemplaron en el modelo.

La regla mencionada se cumple solamente en el modelo con variable psicológica.
Por este motivo elegimos trabajar con el modelo que involucra un factor de riesgo cardiovascular tradicional, un factor psicológico y la edad como factor epidemiológico (modelo de tres factores). Ver tabla 5.

						Intervalo de confianza al 95% para Exp(B)	
	B	Error típico	Wald	Sig.	Exp(B)	Límite inferior	Límite superior
Intersección	-3,043	0,705	18,629	0,000			
E	1,455	0,446	10,628	0,001	4,286	1,787	10,283
DLP	1,585	0,377	17,65	0,000	4,878	2,329	10,218
DR	0,74	0,379	3,802	0,051	2,095	0,996	4,407

Tabla 5. Estimación de los parámetros correspondientes al modelo con los tres factores: E (edad) y DLP (dislipemia), DR (depresión rasgo).

Este último modelo de tres factores ha sido utilizado para construir el predictor. Primero, elegimos una probabilidad de corte p_c y se procedimos de la siguiente forma.

Para el paciente considerado se determinan los valores de las covariables y se calcula p para ese caso, utilizando la fórmula que suministra el modelo de tres factores. Si p resulta mayor (o igual) que p_c se conviene en que el predictor indica la presencia de puntaje de calcio mayor a 400 UH. Si p resulta menor que p_c se conviene que el predictor indica ausencia de puntaje de calcio mayor a 400 UH.

Para cada paciente de la población se comparó con lo observado la predicción para los siguientes valores de la probabilidad de corte: 0,1; 0,2; 0,25; 0,3; 0,4; 0,5; 0,6; 0,7; 0,75; 0,8; 0,9.

A partir de esta comparación se calcularon una especificidad y una sensibilidad para cada probabilidad de corte. Se obtuvieron los siguientes resultados que se agruparon en la tabla 6.

Probabilidad de corte	Sensibilidad	Especificidad
0,1	1	0
0,2	0,95	0,27
0,3	0,84	0,59
0.4	0,84	0,59
0,5	0,63	0,73
0,6	0,63	0,73
0,7	0,17	0,95
0,8	0,17	0,95
0,9	0	1

Tabla 6. Especificidad y sensibilidad para cada probabilidad de corte

Como debe ser, la sensibilidad disminuye mientras que la especificidad aumenta a medida que aumenta la probabilidad de corte.

Para elegir una probabilidad de corte que contemple las exigencias antagónicas de maximizar sensibilidad y especificidad, se construyó un promedio ponderado de ambas: $\varphi_S \cdot S(p_c) + \varphi_E \cdot E(p_c)$

En esta fórmula φ_S es el factor de peso de la sensibilidad y φ_E es el factor de peso de la especificidad. Los pesos son números no negativos cuya suma da siempre uno.

Al asignar un peso distinto a la sensibilidad y a la especificidad se puede asignar mayor importancia a una de éstas, la cual recibe entonces mayor peso. La asignación de pesos puede efectuarse sobre la base de consideraciones clínicas.
La tabla siguiente presenta la correspondencia entre distintos valores de la probabilidad de corte y la correspondiente suma ponderada de sensibilidad y especificidad para factores de ponderación iguales ($\varphi_E = \varphi_S = 0{,}5$).

Probabilidad de corte	Suma ponderada (sensibilidad y especificidad) con factores de ponderación iguales
0,1	0,5
0,2	0,61
0,3	0,715
0.4	0,715
0,5	0,68
0,6	0,68
0,7	0,56
0,8	0,56
0,9	0,5

Tabla 7. Suma ponderada (sensibilidad y especificidad) con factores de ponderación iguales para cada probabilidad de corte

El valor óptimo para la probabilidad de corte cuando los pesos asignados a la sensibilidad y a la especificidad se tomaron iguales, se encuentra comprendido entre 0,3 y 0,4.

Este resultado es el mismo que se obtiene aplicando el índice de Youden $J = Máx_{pc}\{S(p_c)+E(p_c)-1\}$ con el fin de identificar una probabilidad óptima [27].
El índice J de Youden para el predictor de presencia de puntaje de calcio mayor a 400 UH se puede calcular a partir de los datos presentados en la Tabla 7. Se obtiene J=0.43.

Luego, construimos una curva ROC. Para esto, confeccionamos la tabla 8 y el gráfico 1.

Sensibilidad	1-Especificidad
1	1
0,95	0,73
0,84	0,41
0,84	0,41
0,63	0,27
0,63	0,27
0,17	0,05
0,17	0,05
0	0

Tabla 8. Sensibilidad para cada probabilidad de corte y su correspondiente "1-Especificidad"

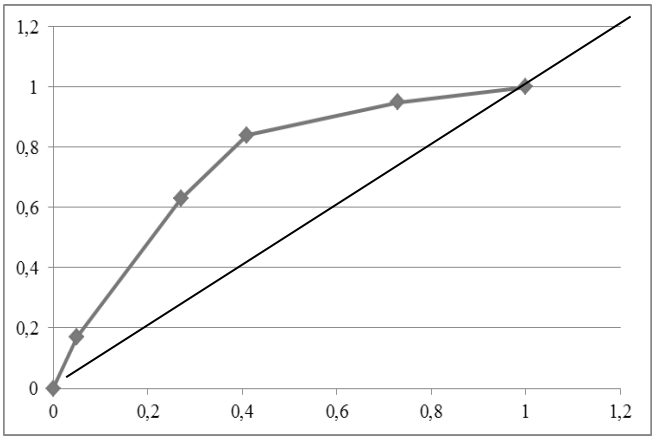

Gráfico 1. Eje x: "1-Especificidad". Eje y: "Sensibilidad"

El área bajo la curva resultó 0,74. Este resultado es estadísticamente significativo ($p<0.05$) frente al valor 0.5 correspondiente a la hipótesis nula (no discriminación). El área bajo la curva ROC suministra una medida del poder de discriminación del modelo construido, en este caso un 74% del máximo posible.

Por último, calculamos los estadísticos "-2 log de la verosimilitud" o -2LL, y comparamos las desviaciones. El mejor ajuste se obtuvo para el modelo de regresión con tres factores (depresión rasgo, dislipemia y edad).

La tabla 9 presenta la desviación obtenida para el modelo ajustado con tres factores (31,402) y la desviación correspondiente al modelo nulo (64,171).

	Criterio de ajuste del modelo
	- 2LL
Modelo nulo	64,171
Modelo ajustado	31,402

Tabla 9. Estadístico "-2 log de la verosimilitud" para el modelo nulo y para el modelo ajustado (modelo final de los tres factores)

La mejora en el ajuste al pasar del modelo nulo al modelo de regresión logística con tres factores es considerable.

Conclusiones de la segunda parte

El modelo simplificado final presenta tres factores en la fórmula que permite calcular la probabilidad de presencia de puntaje de calcio mayor a 400 UH: un factor de riesgo cardiovascular tradicional (dislipemia), un factor psicológico (depresión rasgo) y la edad como factor epidemiológico. Ajusta bien a los datos experimentales.

A partir de la fórmula con tres factores para calcular la probabilidad de presencia de puntaje de calcio mayor a 400 UH, y de un umbral de probabilidad tomado como punto de corte, construimos un predictor.

Comparamos la predicción con el resultado de la exploración tomográfica coronaria y determinamos la variación de la sensibilidad y la especificidad en función de la probabilidad de corte.

El valor del área bajo la curva ROC superó 0,7 lo que permite concluir que el predictor basado en el modelo de los tres factores, cuando se lo ensaya contra la misma población utilizada para construirlo, presenta un poder de descriminación aceptable.

No obstante para validar el predictor de presencia de puntaje de calcio mayor a 400 UH basado en el modelo de regresión es necesario aplicarlo en una población diferente a la empleada para construirlo.

Construir un modelo presenta limitaciones propias de las variables de estudio, del instrumento de medición, del objetivo de estudio y del tamaño de la muestra.

Este trabajo, si bien presenta limitaciones debido al tamaño de la muestra, puede considerarse como un ensayo piloto con méritos suficientes para ser ahondado en otros estudios.

Algunos de los factores de riesgo tradicionales han fallado en la detección de un gran número de individuos en riesgo que luego desarrollaron una patología cardiovascular.

Por tanto, es necesario seguir profundizando sobre la incorporación de otros factores de riesgo no tradicionales.
La construcción de modelos como el que motiva la presente investigación, con un componente psicológico que no había sido incorporado en modelos previos, puede contribuir a la tarea de identificar nuevas interacciones relevantes desde una perspectiva epidemiológica.

Referencias

[1] Bonetti PO, Lerman LO, Lerman A. Endothelial dysfunction: a marker of atherosclerotic risk. Arterioscler.Thromb.Vasc.Biol. 2002; 22(2): 1065–1074.

[2] Lloyd-Jones D, Adams RJ, Brown TM, Carnethon M, Dai S, De Simone G, et al. Executive summary: heart disease and stroke statistics--2010 update: a report from the American Heart Association. Circulation 2010; 121(7):948-954.

[3] Nichols M, Townsend N, Luengo-Fernandez R, Leal J, Gray A, Scarborough P, et al. European Cardiovascular Disease Statistics 2012. European Heart Network and European Society of Cardiology. Sophia Antipolis 2012.

[4] Ross, R. Atherosclerosis an inflammatory disease. The New England Journal of Medicine 1999*;* 340 (2): 115-123.

[5] Lloyd-Jones DM, Larson MG, Beiser A, Levy D. Lifetime risk of developing coronary heart disease. Lancet 1999; 353(9147): 89-92.

[6] Azevedo CF, Rochitte CE, Lima JA. Coronary artery calcium score and coronary computed tomographic angiography for cardiovascular risk stratification. Arq.Bras.Cardiol. 2012; 98(6): 559-568.

[7] Detrano R, Guerci AD, Carr JJ, et al. Coronary calcium as a predictor of coronary events in four racial or ethnic groups. N. Engl. J. Med. 2008; 358(13):1336-45.

[8] Hou ZH, Lu B, Gao, Y, Jiang SL, Wang Y, Li W, et al. Prognostic value of coronary CT angiography and calcium score for major adverse cardiac events in outpatients. JACC Cardiovasc. Imaging 2012; 5 (10): 990-999.

[9] Nucifora G, Bax, J J, van Werkhoven JM, Boogers, M J, Schuijf JD. Coronary artery calcium scoring in cardiovascular risk assessment. Cardiovasc. Ther. 2011; 29(6): 43-53.

[10] Diez Roux AV, Ranjit N, Powell L, Jackson S, Lewis, TT, Shea S et al. Psychosocial factors and coronary calcium in adults without clinical cardiovascular disease. Ann. Intern. Med. 2006; 144(11): 822-831.

[11] Hamer M, Kivimaki M, Lahiri A, Marmot MG, Steptoe A. Persistent cognitive depressive symptoms are associated with coronary artery calcification. Atherosclerosis 2010; 210(1): 209-13.

[12] Janssen I, Powell LH, Matthews KA, Cursio J F, Hollenberg SM, Sutton- Tyrrell, K et al. Depressive symptoms are related to progression of coronary calcium in midlife women: the Study of Women's Health Across the Nation (SWAN) Heart Study. Am. Heart J. 2011; 161 (6): 1186-1191.

[13] Janssen I, Powell LH, Matthews KA, Jasielec MS, Hollenberg SM, Bromberger JT, Sutton-Tyrrell K, Everson-Rose SA. Relation of Persistent Depressive Symptoms to Coronary Artery Calcification in Women Aged 46 to 59 Years. Am. J. Cardiol. 2016; 117(12):1884-9.

[14] O'Malley PG, Jones DL, Feuerstein, IM, Taylor AJ. Lack of correlation between psychological factors and subclinical coronary artery disease. N. Engl. J. Med. 2000; *343*(2): 1298-1304.

[15] Buela-Casal G, Agudelo D, Spielberger CD. IDER: Inventario de Depresión Estado-Rasgo. TEA Ediciones: Madrid. 2008.

[16]Suárez-Bagnasco M. Lesiones ateroscleróticas coronarias, niveles plasmáticos de aldosterona y depresión: un enfoque desde la psico-inmuno-neuro-endocrinología. Tesis de Doctorado en Medicina. Universidad Favaloro. Buenos Aires, 2014.

[17] Barth J, Schumacher M, Herrmann-Lingen, C. Depression as a risk factor for mortality in patients with coronary heart disease: a meta-analysis. Psychosom. Med. 2004; 66: 802-813.

[18] Goldston, K, Baillie AJ. Depression and coronary heart disease: a review of the epidemiological evidence, explanatory mechanisms, and management approaches. Clin. Psychol. Rev. 2008; 28(2), 288–306.

[19] Maes M. Depression is an inflammatory disease, but cell-mediated immune activation is the key component of depression. Prog.Neuropsychopharmacol. Biol. Psychiatry 2011 ; 35(3): 664–675.

[20] Malpas SC. Sympathetic nervous system overactivity and its role in the development of cardiovascular disease. Physiol. Rev. 2010; 90(2): 513–557.

[21] Worster A, Fan J, Ismaila A. Understanding linear and logistic regression analyses. Can.J.Emerg.Med. 2007; 9(2): 111-113.

[22] Wulsin LR Is depression a major risk factor for coronary disease? A systematic review of the epidemiologic evidence. Harv. Rev. Psychiatry. 2004; 12(2): 79-93.

[23] LaValley M. Logistic Regression, Circulation 2008; 117: 2395-2399.

[24] Roca J, Muñoz A. Epidemiologia En: Farreras-Rozman Medicina Interna 24ed. Harcourt: Madrid, 2004.

[25] Bagley S, White H, Golomb B. Logistic regression in the medical literature: Standards for use and reporting with particular attention to one medical domain. Journal of Clinical Epidemiology 2001; 54(10): 979-985.

[26] Hosmer D, Taber S, Lemeshow S. The importance of assessing the fit of logistic regression models: a case study. Am. J. Public Health 1991; 81(12): 1630-1635.

[27] Hosmer DW, Lemeshow S. Applied Logistic Regression. 2nd ed. John Wiley & Sons: New York, 2000.

[28] Kirkwood B, Sterne J. Essential Medical Statistics. Blackwell Science Ltd: Oxford, 2003.

[29] Hemingway H, Marmot M. Psychosocial factors in the etiology and prognosis of coronary heart disease: systematic review of prospective cohort studies. BMJ 1999; 318(7196):1460-1467.

[30] Libby P, Theroux P. Pathophysiology of Coronary Artery Disease. Circulation 2005; 111:3481-3488.

[31] Libby P. Current concepts of the pathogenesis of the acute coronary syndromes. Circulation 2001; 104(3):365–72.

[32] Van der Kooy K, Van Hout H, Marwijk H, Marten, H, Stehouwer C, Beekman A. Depression, and the risk for cardiovascular diseases: systematic review and meta-analysis. Int.J. Geriatr. Psychiatry 2007; 22(7): 613-626.

[33] Perduzzi P, Concato J, Kemper E, Holford T, Feinstein A. A simulation study of the number of events per variable in logistic regression analysis. J.Clin.Epidemiol. 1996; 49(12):1273-1379.